



大数据及其关键技术的教育应用实证分析*

胡水星

(湖州师范学院 教师教育学院, 浙江湖州 313000)

[摘要] 大数据作为信息技术发展的新趋势,已经渗透到各行各业,成为重要的生成因素。在数据分析驱动教育,变革教学的大数据时代,教育领域同样蕴藏着具有广泛应用价值的海量数据,在对教育数据挖掘和学习分析技术等教育大数据关键技术分析的基础上,结合共词分析和教育博客等社会化网络教育数据,构建教育领域的相关学习分析和数据挖掘模型,探索教育变量之间的相关关系,实践大数据的教育应用,为教育教学提供有效的决策支持服务,促进教育教学的变革与创新。

[关键词] 大数据; 学习分析; 数据挖掘; 共词分析; 社会网络分析

[中图分类号] G434 [文献标识码] A [文章编号] 1672-0008(2015)05-0046-08

DOI:10.15881/j.cnki.cn33-1304/g4.2015.05.006

大数据作为信息技术发展的新趋势,已经渗透到各行各业,成为重要的驱动因素,并掀起行业变革的巨浪。随着我国教育信息化水平的不断提升,越来越多的学习管理系统在教育领域中被应用,数据化的学习信息和学生信息逐渐增多,教育数据的海量增长,导致在教育管理、教育服务、教学研究、教育评价等领域也面临大数据问题,教育大数据时代已经悄然来临。“教育信息量的爆炸式增长以及相关数据处理技术的创新发展成为人们新的视野焦点,如何利用与分析这些数据信息,不仅影响着信息交流、知识传递和学习效果,更在一定程度上影响着教学决策制定和学习模式优化,已经成为目前教育工作者和学习者一致关注的方面。”^[1]

在数据分析驱动教育,变革教学的大数据时代,教育领域同样蕴藏着具有广泛应用价值的海量数据,利用教育数据挖掘和学习分析技术,构建教育领域的相关模型,探索教育变量之间的相关关系,为教育教学决策提供有效的支持将成为未来教育的发展趋势。

一、大数据及教育大数据关键技术

计算机技术、移动通信技术、互联网技术,再到

大数据技术的登台亮相,不仅体现着信息技术的深入与发展,更体现着大数据时代的到来;大数据技术在教育领域的广泛应用引起了教育变革,推动了教育发展,学习分析实现了教育的探索性分析,数据挖掘实现了教育管理与决策的制定。学习分析和数据挖掘这两大数据关键技术在教育领域的深入应用,为教育教学带来了新的机遇。

(一)大数据

2012年,联合国发布了大数据白皮书“*Big Data for Development: Challenges & Opportunities*”,明确提出大数据时代已经到来。^[2]大数据作为信息技术发展的新趋势,具有海量的数据规模(Volume)、快速的数据流转(Velocity)、多样的数据类型(Variety)和巨大的数据价值(Value)4V特性。^[3]大数据是一种价值观、方法论……是一场思维的大变革,已经渗透到各行各业,成为重要的一种生成因素,并掀起了变革的浪潮。^[4]通过对海量数据的分析挖掘,以一种前所未有的方式获得巨大的产品服务、深刻的真知灼见,为我们理解生活以及认识世界提供了一种全新的思维方式,实现思维的三大转变:一是不再依赖于小样本数据,而是与现象相关的所有数据;二是不再热衷于追求微观层面的精确,而是宏观层面的洞察力;三是从

*基金项目:本文系全国教育科学“十二五”规划2014年度教育部重点课题“现代服务业视野下的社区教育信息化构建与实践研究”(课题批准号:DCA140232)的研究成果。

传统的因果关系追求中解脱出来,关注相关关系的发现和应用。^[5]

大数据借助无所不在的传感设备和计算能力,对现实世界、虚拟世界以及虚实融合世界的复杂网络数据进行解析和挖掘,实现行为判断和决策。随着教育领域各种学习管理系统、课程管理系统、网络互动平台的应用,各种学习行为、学习状态等教育数据将海量增长,教育大数据时代已经悄然来临。大数据使得教育信息成为可捕捉、可量化、可传递的数字存在,大数据使得教育过程性考察成为可能,更能透过真实的数据发现教与学的关系,教育正悄悄地发生着一场革命。^[6]

(二)教育大数据关键技术

近年来,随着大数据的推进与发展,教育大数据处理与分析已经成为推动教育改革与发展的驱动力,引起了各国政府和教育行政部门的高度重视。2012年10月美国教育部发布了《通过教育数据挖掘和学习分析促进教与学》(Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics)报告。^[7]力图通过教育大数据分析挖掘,促进“大数据”教育应用,从而为教育发展抢得先机。

大数据为海量教育数据的存取提供了技术基础,但原始的教育数据只是教育大数据的基础,只有通过采集到的各种数据进行教育数据挖掘,构建学习分析模型,发现教育变量之间的关系,并实现赋予数据相关意义,才能使数据转变为信息;信息进一步经过分析和综合,形成知识;最后通过实践运用,知识才上升到智慧层次。^[8]因此,教育数据挖掘和学习分析技术是教育大数据的关键技术。

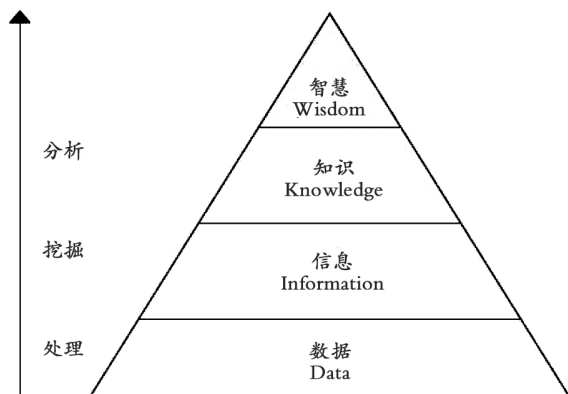


图1 数据演变过程

1. 教育数据挖掘

教育数据挖掘 (Education Data Mining) 是综合

运用数理统计、人工智能与机器学习和数据挖掘等技术与方法,对教育原始数据进行分析处理,通过构建数据模型,对学习者的学习结果与学习内容、学习资源和教学行为等变量进行相关关系分析,从而有效地预测学习者未来的学习趋势。^[9]并为教育工作者、学习者、学生家长、教育教学研究者以及教学软件开发人员提供支持,实现教育系统中教育资源的良性互动,最终实现改进学习的目的。教育数据挖掘流程具体如图2所示。^[10]

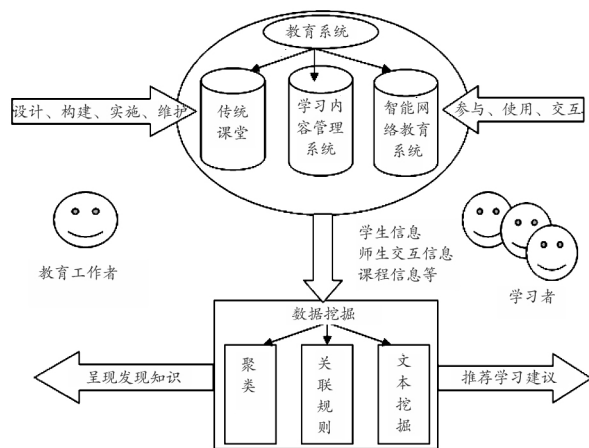


图2 教育数据挖掘流程

2. 学习分析

学习分析 (Learning Analytics) 的研究对象是学生及其学习环境,目的是通过对教育海量数据的分析和建模,发现潜在问题,优化和理解学习,预测学习者在学习中的进步和表现。学习分析具体过程如图3所示。显而易见,学习分析技术就是围绕学习者在学习过程中产生的各种信息数据,利用各种数学建模方法和数据处理技术来解释这些数据,并根据结果数据与分析信息,对学习过程与情境进行探究,并进一步发现学习的规律,为优化和完善教学提供相应的反馈,持续地促进学习者的学习。学习分析技术利用数据挖掘、数据解释与数据建模的优势,对学习平台中积累的大量数据信息进行采集、存储、分析和表示,并运用分析测量结果对学习者的学习行为进行评估、预测和干预,为个别学生量身定制更有效的教育,进而改善和提升教与学的质量与效能,实现改善教学和促进学习的目的。^[11]

除了学习分析和数据挖掘这两大教育大数据关键技术,近年来慕课、微课、翻转课堂、社会网络软件、云计算、Moodle等网络学习开源平台、WEB2.0技术都可以纳入教育大数据的技术范畴。

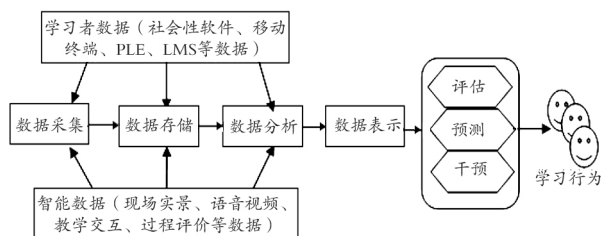


图3 学习分析过程模式

二、大数据促进教育变革

随着教育领域各种大数据技术的应用,我们已经进入一个数据驱动教学、分析变革教育的大数据时代。虽然技术还没有给教育带来如1987年Alfred Bork所预言那样的变革^[12],但技术具有拓展、增强或者变革人类学习的潜力,科技进步不可避免地带来了教育的革新,并且对教育发展带来了深远的影响。如,大规模数据分布式处理技术实现了结构化与非结构化数据存储;云计算技术实现了教育资源的共建共享;社会性交互软件促进了学习型社会的构建;MOOCs与翻转课堂实现了个性化学习构建;开源学习管理系统实现了学习过程管理。大数据技术在教育领域的应用革新了教育思维方式、重构了教学评价方式、颠覆了传统教学模式、实现了个性化教育,教育大数据背景下教学的规训与教化在撤退,支持和服务在推进。^[13]

(一)数据挖掘与学习分析有助于教学决策和评价

教育大数据记录了教学的过程,发现了新的知识,创造了更大的教育价值,促进和优化了教学策略和评价。移动互联网时代,知识的获取变得以学生为中心,因为,每个学生的智力特点和吸收水平都是不一样的,移动互联网支持了以学生为中心的学习,突破传统不再是所有人在统一的课堂上在规定的四十五分钟内听相同的教学内容。新兴的教育技术与资源使得教育更加以学习者为中心,使教育从批量到个性的实现成为一种可能;教师的教育思维也从宏观的群体教育向微观的个体教育方式转变,促进了以学习者为中心的个性化教育的实施,进一步使得因材施教成为一种可能。

从技术层面上说,学习者在互联网等媒体上留下的任何数据痕迹,都可以进行分析,可以发现数据背后隐藏的学习者相关学习特征、兴趣爱好、行为倾向,与教育教学相关的状态信息都将一览无余。从这个意义上来说,未来的教育发展方向就是应用学习

分析和数据挖掘等大数据技术去实现精准的个性化教育。将大数据对教育教学中海量数据的整合分析,结合态度、行为和背景等因素,就可以发现学生思想、行为和心态的变化,分析出每个学生的特点,再结合总体学生的表现和其他因素的分布就可以准确对某位学生实施正确的评价,最终使得对每个学生的教学都可以建立在对过去行为数据的分析基础上。^[14]从而真正打破传统技术背景下“不得不承认,对于学生,我们知道的太少”那样的窘境。^[15]

大数据技术有利于对教师课堂教学计划、课堂教学评价、课堂视频资源中的各种数据进行提取和分析,从而为预测、处理教学行为、学习心理提供了重要依据,为教学评价提供了较为全面和完整的信息,克服了评估主观性强的缺憾,教学评价不再是经验式的,而是在大数据基础上的“归纳”,更具说服力和公信力,实现了教学评价的客观公正与科学正确,教学决策的针对性与时效性。

(二)微课与翻转课堂教学有利于个性化学习模式重构

未来教育在大数据技术的支持下变得越来越个性化,慕课、微课与翻转课堂的教学应用有利于个性化学习环境的构建,数字化课程资源的标准化定制实现了学习内容的学生自组织学习,在线学习使得学校教育和教师更多地转向学生个体,关注学习者的个性化培养,教师实现了从教学者到助学者角色的转变,更多地是承担学习的支持服务和协作交流。^[16]个性化学习模式更多地关注师生之间、生生之间、学生与教育媒体之间的交往互动、个性化服务和灵活的教学范式。微课实现了知识从固化到碎片,移动互联网则使我们可以充分利用碎片化的时间。不仅如此,我们每一次对碎片化时间的利用都还可以是非常高效的,因为,学习系统了解你的学习情况,知道怎么让你更高效地学习。

传统教育是大家在一个封闭的客厅里,规定每节课45分钟,而现在移动互联网可以让学生和全世界各个地方的学习者交流,做到线下进行知识学习,线上进行知识传递,完全颠倒传统的教学结构,使强调知识传递、以教定学的知识传授模式逐步让位于强调问题中心、以学为主的整合探究模式,构建了新的教学范式。

教师可以利用微课和翻转课堂进行基于技术的课堂教学案例欣赏,积极开展分析、讨论和教学反思活动,在学习活动中实现对技术、学科和教学法之间

的深层次理解。^[17]在这种模式下,一方面,充分调动了学生的积极主动性,学习者在课堂上进行深入交互和思想碰撞,而不再是存在了几个世纪的那种被动的教学模式接受者;另一方面可以进一步加深对整合技术的学科教学知识的理解,培养教师将技术和学科教学知识整合地根植于自身课堂中的能力,实现将知识的传授迁移到了课外。

(三)在线视频与大规模开放课程有利于教学知识呈现

开放存取已经成为帮助所有人进行学习的关键因素之一,“技术支持了信息的开放存取,实现了知识共享的无处不在”,^[18]这也是21世纪由新的学习技术所带来的张力和机会:学科内容的音频、视频等多媒体融入呈现,营造了更为丰富的交流互动和学生体验参与,促进了学习者情感领域的认知与发展,并进一步帮助我们所有人学习;现有的视频动画、网络多媒体、甚至3D的教学内容,生动逼真地呈现在学习者面前,使学习突破了传统的学校围墙的限制,未来的师生关系应该是从传道授业变成解惑的角色,传道授业由互联网上高质量的视频资源来替代;在线视频分享为每一位学习者打开了一扇从其他机构那里学习的大门,学习和培训从单中心到多中心,各种大规模开放在线课程让学习方式越来越丰富。通过向学习者提供观看、参与和反思的教学视频资源,实现了跨界学习交互和教学应用示范。这种交互和示范不仅能够展示如何理解技术、学科内容知识和学科教学知识之间的关系,更能够展示如何为学科内容领域的教学制定与技术有关的有效决策,^[19]从而促进学习者自主建构知识体系,实现教学知识的呈现与内化。

(四)学习管理系统和web2.0技术有助于教学管理和情境参与

利用web2.0技术,用户可以在其中创建视频群组,订阅特定频道和成员的视频,上传视频,给视频贴标签。在世界范围内分享视频、创建播放列表、与拥有共同兴趣的人保持联系等。这些工具赋予了学习者在教育过程中的话语权,支持参与式学习。^[20]大数据促进教育发展的第三大宏观趋势是建构一种协同、协商和共享这些知识和信息的文化创造;一种参与式学习文化。^[21]正如约翰·赛利·布朗所指出的那样,在这种新型的参与式教育环境中,学习者开始参与到一种建构、改进、学习、分享的文化中。各种免费的教育资源随处可见、各种支持学习者参与的社会

性工具软件层出不穷,资源和工具的结合能使学习者自由添加、实时评论和独立创建全新资源,这时我们需要重新审视和界定学习究竟是什么——学习不再是消费和吸收,而变成了生成和参与。由于有了Moodle、Sakai、Blackboard等学习管理系统和web2.0技术对学习参与的支持,实现了以学习需求和学习者为中心的转变,动态的教与学关系,通过挖掘知识、寻求联系、总结规律,每个教师和学习者都能发挥自己最大教育潜能。

三、大数据关键技术教育应用实证分析

大数据所带来的教育的未来,不单是表述一个理念和社会直觉,更是一种教育变革和社会行动。在大数据技术的支持下,教育不再是单一的理念交流与经验传承的社会科学,而确实变成依托教育行为数据的实证科学。在教育大数据时代,由于教学实验的布置、教育时空的设计、学习场景的构建、教育环境的创设、教育数据的采集和教育管理的决策,这些过去只能依靠理念灵感加经验的东西,可以实实在在地以数据形式进行记录、跟踪,真正变成一种数据支撑的行为科学。^[22]

大数据技术在教育教学实践中的广泛应用,积累了大量的学习状态数据,如何让数据说话,揭示教育数据背后的教育现象本质,为教育决策和政策制定提供建设性解决方案,我们将通过数据挖掘和学习分析这两大数据关键技术在教学领域的应用案例进行实证分析,以期数据挖掘和学习分析等大数据技术的教育应用提供借鉴与启发。

(一)基于教育数据挖掘的共词分析

教育大数据技术促进了教师专业发展,两者具有一定的相关性,国内许多文献对两者进行了深入研究,我们抽取10年以来在CNKI、EBSCO、Springer link等国内外期刊上出现的教师专业发展和教育大数据相关的关键词频次,利用内容分析、文本分析等方法对关键词出现的共篇关系构建共词分析矩阵,并通过Ochia相似性系数法对数据进行分析计算,从而将共词矩阵转换为相关矩阵,在相关矩阵的基础上,进一步构建相异矩阵,并采用多维度尺度分析法进行分析,形成教育大数据背景下教师专业发展的知识体系谱系图,从而对教育大数据教师应用提供良好的决策指导。

期刊论文一般都有几个关键词,用以说明本篇论文的主要研究领域和重点方向等,如果多个关键

词在多篇期刊论文中重复出现,这说明这些关键词之间存在某种关系;如果几个关键词共现的频次比较大,说明这些关键词关系稳定,可以看成一种必然的联系。

利用文献管理软件 NoteExpress 对大数据和教师专业发展相关关键词进行统计分析,并按关键词出现的频次进行分类汇总、统计分析和数量排序。为了数据挖掘过程的清晰展现,提取重复出现频次较高的 10 个关键词为高频关键词,以内容分析的视角建立高频关键词共词矩阵,具体如表 1 所示。

表 1 高频关键词共词矩阵

共现 词条	大数 据	数据 挖掘	学习 分析	翻转 课堂	慕课	个性 化学习	教师 专业 发展	教师 TPA CK	教师 信息 素养	教学 绩效 评估
大数据	124	26	20	12	22	11	12	15	12	25
数据挖掘	26	122	29	11	13	10	6	35	15	20
学习分析	20	29	120	1	0	4	2	25	15	30
翻转课堂	12	11	1	85	1	3	7	15	20	10
慕课	22	13	0	1	59	16	1	18	25	12
个性化学习	11	10	4	3	16	54	8	7	8	8
教师专业发展	12	6	2	7	1	8	51	50	30	15
教师 TPACK	15	35	25	15	18	7	50	70	45	20
教师信息素养	12	15	15	20	25	8	30	45	50	18
教学绩效评估	25	20	30	10	12	8	15	20	18	45

共词矩阵是一个相关关系矩阵,其对角线上的数据表示为某关键词出现的频次,如关键词“大数据”共出现 124 次,它与数据挖掘同时出现在 26 篇期刊文献中,也可以理解为在 26 篇期刊论文的关键词中上述两个关键词共同出现。

为了对共词矩阵数据进行多元统计分析,挖掘其隐含的知识信息,需要对矩阵的数据结构进行变换,借鉴 Ochiai 相似性系数法则,实现对共词矩阵进行相似矩阵的转换。Ochiai 相似性系数法的具体计算公式为: $Ochiai \text{ 系数} = N_{ij} / (N_i * N_j)^{1/2}$ 。其中 N_i 和 N_j 分别代表关键词 i 和 j 出现的频次, N_{ij} 表示两者共同出现的频次。经过统计分析,我们可以得到如表 2 所示的。

表 2 表征大数据与教师专业发展相关的关键词相似矩阵

共现 词条	大数 据	数据 挖掘	学习 分析	翻转 课堂	慕课	个性 化学习	教师 专业 发展	教师 TPA CK	教师 信息 素养	教学 绩效 评估
大数据	1	0.2114	0.1640	0.1169	0.2572	0.1344	0.1509	0.1610	0.1524	0.3347
数据挖掘	0.2114	1	0.2397	0.1080	0.1532	0.1232	0.0761	0.3787	0.1921	0.2699
学习分析	0.1640	0.2397	1	0.0099	0	0.0497	0.0256	0.2728	0.1936	0.4082
翻转课堂	0.1169	0.1080	0.0099	1	0.0141	0.0443	0.1063	0.1945	0.3068	0.1617
慕课	0.2572	0.1532	0	0.0141	1	0.2835	0.0182	0.2801	0.4603	0.2329
个性化学习	0.1344	0.1232	0.0497	0.0443	0.2835	1	0.1524	0.1139	0.1540	0.1623
教师专业发展	0.1509	0.0761	0.0256	0.1063	0.0182	0.1524	1	0.8368	0.5941	0.3131
教师 TPACK	0.1610	0.3787	0.2728	0.1945	0.2801	0.1139	0.8368	1	0.7606	0.3563
教师信息素养	0.1524	0.1921	0.1936	0.3068	0.4603	0.1540	0.5941	0.7606	1	0.3795
教学绩效评估	0.3347	0.2699	0.4082	0.1617	0.2329	0.1623	0.3131	0.3563	0.3795	1

相似矩阵中的数字也称为相似数据,其大小表明相应的两个关键词之间距离远近的程度。为了方便进一步进行多维尺度分析,可以用 1 与全部相似矩阵上的数据相减,得到相异矩阵,其数字大小表示两关键词间相异程度,具体如表 3 所示。

表 3 表征大数据与教师专业发展相关的关键词相异矩阵

共现 词条	大数 据	数据 挖掘	学习 分析	翻转 课堂	慕课	个性 化学习	教师 专业 发展	教师 TPA CK	教师 信息 素养	教学 绩效 评估
大数据	0	0.7886	0.836	0.8831	0.7428	0.8656	0.8491	0.839	0.8476	0.6653
数据挖掘	0.7886	0	0.7603	0.892	0.8468	0.8768	0.9239	0.6213	0.8079	0.7301
学习分析	0.836	0.7603	0	0.9901	1	0.9503	0.9744	0.7272	0.8064	0.5918
翻转课堂	0.8831	0.892	0.9901	0	0.9859	0.9557	0.8937	0.8055	0.6932	0.8383
慕课	0.7428	0.8468	1	0.9859	0	0.7165	0.9818	0.7199	0.5397	0.7671
个性化学习	0.8656	0.8768	0.9503	0.9557	0.7165	0	0.8476	0.8861	0.846	0.8377
教师专业发展	0.8491	0.9239	0.9744	0.8937	0.9818	0.8476	0	0.1632	0.4059	0.6869
教师 TPACK	0.839	0.6213	0.7272	0.8055	0.7199	0.8861	0.1632	0	0.2394	0.6437
教师信息素养	0.8476	0.8079	0.8064	0.6932	0.5397	0.846	0.4059	0.2394	0	0.6205
教学绩效评估	0.6653	0.7301	0.5918	0.8383	0.7671	0.8377	0.6869	0.6437	0.6205	0

相异矩阵中的数据也可称之为不相似数据,其数值的大小表明关键词之间相异的程度,如数据越大则表明关键词之间相似度越小,距离越远;反之则表明关键词之间的相似度越大,距离越近。^[14]基于上述的相异矩阵,我们采用 SPSS 的分析—度量—多维尺度进行关键词图谱分析,得到如图 4 所示的教育大数据与教师专业发展高频关键词多维尺度图。

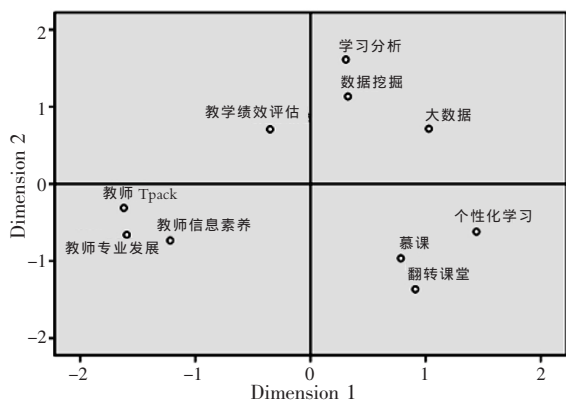


图4 教育大数据与教师专业发展高频关键词多维尺度图

从上面的“教育大数据与教师专业发展高频关键词多维尺度图”中可以发现象限一、象限三和象限四是几块相对集中的研究领域。其中第一象限属于大数据范畴,主要关注“大数据”、“学习分析”和“数据挖掘”;第三象限主要属于教师专业发展范畴,主要包括“教师信息素养”、“教师 Tpack”、“教师专业发展”;第四个象限属于个性化学习技术范畴,关注“慕课”、“翻转课堂”和“个性化学习”;相对这几个关键词集中象限,第二象限只有教学绩效评估,这说明教学绩效评估和其他几个范畴共同联系相对少些。关键词的图中分布也比较符合我们对相关文本内容的分析,对教师教育专业化发展具有很好的决策指导作用,如,在进行大数据学习应用时应该重视学习分析和数据挖掘的学习;在进行教师专业化培训时应该重视教师 Tpack 能力和教师信息素养的培养;在进行个性化学习支持服务的时候应积极利用翻转课堂和慕课进行学习环境构建、学习模式创设。

通过教育数据挖掘共词分析,很好地展现了教育大数据背景下教师专业发展的知识体系谱系图,更加明确地解释出教育大数据对教师专业发展的影响及两者之间的相关性,从而为教育大数据背景下的教师专业发展提供确实有效的指导与建议。

(二)基于社会网络分析的教育博客学习分析

社会网络分析可以对网络群体中的个体学习行为与学习行动者之间的关联和群体影响进行分析,通过相关的社会网络分析软件,利用距离、密度和中心度、社群图等概念对网络属性进行评价,是社会结构关系研究可视化的、说服力强的一种方法手段。教育大数据视野下,学习分析技术越来越被重视,社会网络分析作为一种学习分析方法,可以对网络学习平台、课程管理平台、网络日志等记录的大量过程化数据进行深入挖掘,发现网络学习中学习者的学习

行为和社会交互特征,从而为网络学习提供决策引导和服务支持。

为了运用社会网络分析进行教育数据学习分析,我们选择一个浙江省中小学教师专业发展培训班作为样本数据,依托湖州师范学院精品课程《教育技术学基础》为网络平台,具体如图5所示。通过对培训教师在该课程网站上的教育博客群参与活动分析,选择培训班20名教师的教育博客交流情况,进行网络日志数据的收集。并规定博客的发帖、链接、回复和推荐都记为一次网络交流。



图5 “教育技术学基础”教育 blog 平台

数据的处理分析主要采用社会网络分析软件 UCINET6.0,数据挖掘以前,先根据教育博客上的网络日志进行数据获取,然后对参与博客交流的教师进行 T1-T20 的编码,培训教师用 TEC 编码标示。关系矩阵的建立,具体如表4所示。

表4 现代教育技术培训教师参与博客关系矩阵

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20	TEC
T1	1	2	1	1	2	2	3	1	1	2	3	4	2	1	2	3	0	0	1	2	3
T2	1	1	2	1	2	3	1	1	1	0	1	2	3	1	0	1	1	1	1	2	2
T3	1	0	1	1	1	1	2	1	2	1	3	2	3	4	5	1	0	1	2	3	4
T4	1	1	1	1	1	1	1	1	3	2	2	2	5	1	1	1	2	3	4	1	3
T5	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	1	2	3
T6	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	2	1	1	0	1	3
T7	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	3
T8	0	0	0	0	0	1	1	1	1	1	1	1	1	1	2	2	2	2	2	1	3
T9	1	2	2	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	2	2	3
T10	1	1	1	1	1	1	1	2	0	1	0	1	1	2	3	1	2	1	2	2	4
T11	1	2	0	1	0	1	0	1	1	1	1	0	0	0	0	0	1	1	2	3	3
T12	1	2	3	4	1	1	0	0	0	0	1	1	1	0	0	0	0	2	2	3	2
T13	1	0	0	0	2	2	2	2	2	2	3	3	1	3	2	2	2	1	1	1	5
T14	1	0	2	2	2	2	2	1	1	2	2	2	1	1	0	0	1	1	1	1	6
T15	3	1	1	2	3	2	2	2	2	2	3	3	1	0	1	0	0	2	2	1	8
T16	1	1	1	0	1	0	1	1	1	2	2	2	2	1	0	1	0	1	2	2	7

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20	TEC
T16	1	1	1	0	1	0	1	1		1	2	2	2	1	0	1	0	1	2	2	7
T17	1	1	1	1	1	1	1	2	0	1	0	1	1	1	1	1	1	1	2	2	5
T18	1	2	0	1	0	1	0	1	1	1	1	0	1	2	1	2	1	1	1	2	5
T19	1	2	3	4	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	5
T20	1	0	0	0	2	2	2	2	2	2	3	3	2	3	1	1	1	1	0	1	9
TEC	2	2	2	3	3	3	2	3	3	5	2	2	5	6	8	7	5	5	5	8	1

行列交叉单元格中的数据表示第*i*行的教师对第*j*列的教师进行教育博客互动的次数,0表示没有互动,每个教师相对于本人,博客互动都定为1。经过 Netdraw 导入关系矩阵,经过 UCINET6.0 教师教育博客互动交流社群图分析,具体如图 6 所示。

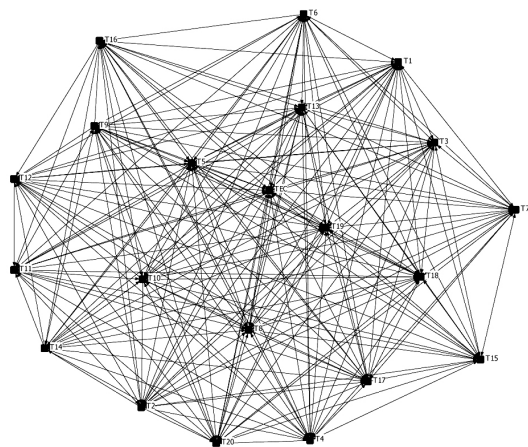


图 6 教师教育博客互动交流社群图

从社群图中可以看出,TEC 在网络博客群中处于中心地位,而且没有孤立的点,这说明教师专业化培训中培训教师处于主导地位,和受培训教师积极进行网络博客交流与互动,帮助受训教师解决问题;同时受训教师也积极相互进行合作交流,进行教育技术课程学习。

进一步通过 networks—centrality—degree 进行中心度分析,得到如表 5 所示。

从表 5 中可以看出,TEC 的入度和出度都是最高的,这进一步说明了教师在培训中的地位,起到了协作、引导和帮助的作用;同时 T15 出度比较多,说明该教师在培训中能够通过积极发帖引导其他教师进行培训交流;T20 的入度比较多,说明其他参与培训的老师比较喜欢和该教师进行互动交流,同时也表明该教师在教育博客中的声望较好,比较受广大培训教师的欢迎。

为进一步了解各培训教师在网络中的交互特征和作用,我们进一步通过 networks—centrality—freeman betweenness—node betweenness 进行中间中心度分析,得到表 6 所示。

表 5 中心度分析表

序号	教师编号	OutDegree	InDegree	NrmOutDeg	NrmInDeg
21	TEC	81.000	86.000	45.000	47.778
15	T15	40.000	32.000	22.222	17.778
3	T3	38.000	24.000	21.111	13.333
4	T4	37.000	28.000	20.556	15.556
20	T20	37.000	40.000	20.556	22.222
1	T1	36.000	22.000	20.000	12.222
13	T13	36.000	34.000	20.000	18.889
9	T9	32.000	23.000	17.778	12.778
5	T5	32.000	27.000	17.778	15.000
14	T14	30.000	32.000	16.667	17.778
10	T10	28.000	27.000	15.556	15.000
2	T2	27.000	23.000	15.000	12.778
16	T16	26.000	29.000	14.444	16.111
17	T17	25.000	24.000	13.889	13.333
19	T19	25.000	34.000	13.889	18.889
6	T6	25.000	29.000	13.889	16.111
18	T18	24.000	28.000	13.333	15.556
12	T12	23.000	34.000	12.778	18.889
8	T8	22.000	26.000	12.222	14.444
11	T11	18.000	33.000	10.000	18.333
7	T7	17.000	24.000	9.444	13.333

表 6 中间中心度分析

序号	教师编号	Betweenness	nBetweenness
21	TEC	5.026	1.323
6	T6	4.496	1.183
1	T1	4.097	1.078
4	T4	3.693	0.972
5	T5	3.589	0.944
20	T20	3.410	0.897
18	T18	3.310	0.871
19	T19	3.293	0.867
9	T9	3.195	0.841
2	T2	3.074	0.809
10	T10	3.029	0.797
17	T17	2.943	0.774
13	T13	2.900	0.763
14	T14	2.628	0.691
3	T3	2.519	0.663
8	T8	2.415	0.636
15	T15	2.404	0.633
7	T7	2.290	0.603
11	T11	2.142	0.564
16	T16	2.013	0.530
12	T12	1.534	0.404

从表 6 中可以发现 TEC、T6 和 T1 的中间中心度比较高,这说明教师和这两位培训学员在博客讨论区中扮演了领导者的角色,对教学培训资源具有一定的控制作用;同时这两位受训教师也起到了培训的桥梁沟通作用,对教师培训交流的信息起到了很好地沟通传递作用。

四、结语

大数据作为信息技术发展的新趋势,已经渗透到各种领域,成为变革的一种重要生成因素。在数据分析驱动教育,变革教学的大数据时代,教育领域同样蕴藏着具有广泛应用价值的海量数据。多元的教

育“数据”已是教育行业及其相关领域必须考虑的因素。如何高效分析和挖掘这类数据信息已经成为提高教育质量,推动教育发展的新浪潮。大数据关键技术正是从技术层面实现了各种学习行为数据的量化与显现,使学习者的体验感受得以实现,通过记录、分类、挖掘和运用学生学习过程中产生的大量非结构化数据,能够揭示出学生的学习模式;通过监测、跟踪、分析和应用学生在整个学习过程中形成的数据档案,能够帮助理解学生为了掌握学习内容而进行学习的全过程。^[23]

学习分析和数据挖掘这两大数据关键技术已经成为教育研究的热点和趋势,通过数据挖掘可以对教育数据进行从微观到宏观的统计、分析、综合和推理,指导教育教学实际问题的解决,发现教育现象之间的相互关联和规则,从而可以更好地做出教育预测和实施教育决策;通过学习分析既能为学生提供高质量、个性化的学习体验,又能改进教育工作者的教学方式,完善和优化教学过程,真正使教学实践活动转向关注微观个体,实现真正意义上的个性化学习和教育个性化支持服务,从而有利于促进教育公平。

总之,学习分析和数据挖掘等大数据关键技术的教育应用实践,不仅探索了教育教学数据背后隐藏的相关关系,更为教育教学决策的科学制定提供了有效的服务支持,促进了教育教学的变革与创新,引领了教育教学发展的新方向。

[参考文献]

- [1]周馨.大数据时代教育数据价值挖掘[J].信息与电脑(理论版),2013,(8).
[2][3][7][9]徐鹏,王以宁,刘艳华,张海.大数据视角分析学习变革——美

国《通过教育数据挖掘和学习分析促进教与学》报告解读及启示[J].远程教育杂志,2013,(6).

- [4]张杰夫.大数据—大视野—大教育[J].中小学信息技术教育,2013,(10).
[5]维克托·迈尔·舍恩伯格,肯尼思·库克耶著.盛杨燕,周涛译.大数据时代[M].杭州:浙江人民出版社,2013,(1).
[6]魏忠.教育正悄悄发生一场革命[M].华东师范大学出版社,2014,(11).
[8]马婧,韩锡斌,周潜,程建钢.基于学习分析的高校师生在线教学群体行为的实证研究[J].电化教育研究,2014,(2).
[10]魏顺平.教育数据的挖掘、分析、应用[J].中小学信息技术教育,2013,(10).
[11]杨庆安,赵伟男,张海.大数据在教育领域应用的学习分析框架构建[J].软件导刊·教育技术,2013,(5).
[12]Bork, A. Learning with personal computers [M]. New York: Harper & Row, 1987.
[13]魏忠.教育正悄悄发生一场革命[M].上海:华东师范大学出版社,2014,(11).
[14]喻长志.大数据时代教育的可能转向[J].江淮论坛,2013,(7).
[15]张燕南,赵中建.大数据时代思维方式对教育的启示[J].教育发展研究,2013,(21):3-4.
[16]魏忠.大数据时代的教育革命[EB/OL].[2014-07-29]. http://www.edu.cn/bigdata_12674/20140729/t20140729_1157265.shtml.
[17]闫志明,徐福荫. TPack: 信息时代教师专业化的知识基础[J].现代教育技术,2013,(3).
[18]黄德群.开放的学习世界:网络技术如何变革教育[J].韶关学院学报,2012,(12).
[19]全美教师教育学院协会创新与技术委员会主编,任友群,詹艺主译.整合技术的学科教学知识:教育者手册[M].北京:教育科学出版社,2011.
[20][21]柯蒂斯·邦克著,焦建利译.世界是开放的——网络技术如何变革教育[M].上海:华东师范大学出版社,2010.
[22]魏忠,张芳芳.综述:大数据与教育革命.[EB/OL] http://blog.sina.com.cn/s/blog_537ef17301011qe9.html.
[23]胡德维.大数据“革命”教育[N].光明日报,2013-10-19.

[作者简介]

胡水星,湖州师范学院教师教育学院副教授、博士,研究方向:教育信息化、教育数据挖掘与学习分析。

An Empirical Analysis of Big Data and Its Key Technology in Educational Practice

Hu Shuixing

(School of Teachers Education, Huzhou Teacher's University, Huzhou, Zhejiang 313000)

[Abstract] As a new trend of the development of information technology, Big data has become an important factor in different trades. In the age when data analysis driving education and transforming teaching, there have big data with widespread application value in the field of education. Based on utilizing the key technology about data mining and learning analytics, combining with co-word analysis in education data and social networking blog, a model of learning analytics and data mining in the field of education is constructed. This model can explore the relationship between educational variables, be applied in education, provide an effective decision support service for teaching, and promote the reform and innovation of education.

[Keywords] Big data; Learning analytics; Data mining; Co-word analysis; Social network analysis

收稿日期:2015年7月18日

责任编辑:黄晓磊 吕东东